Faculty Working Papers

MULTIDIMENSIONAL SCALING OF BINARY DATA
FOR HOMOGENEOUS GROUPS

Robert P. Redinger and Jagdish N. Sheth

#411

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

June 29, 1977

MULTIDIMENSIONAL SCALING OF BINARY DATA FOR HOMOGENEOUS GROUPS

Robert P. Redinger and Jagdish N. Sheth

#411

MULTIDIMENSIONAL SCALING OF BINARY DATA FOR HOMOGENEOUS GROUPS

Robert P. Redinger
and
Jagdish N. Sheth
University of Illinois

ABSTRACT

A new methodology is proposed for perceptual mapping as an alternative to the nonmetric multidimensional scaling. The new methodology requires binary similarity judgments and utilizes the Ekart-Young decomposition procedure for mapping objects on a multidimensional space.

The new technique is tested with respect to mapping of fifteen brands of soft drinks. The same study also collected rank-ordered similarity judgments and utilized the standard nonmetric scaling techniques as a comparative test. As expected, binary judgments were more reliable and produced a more meaningful multidimensional space than the nonmetric procedure.

MULTIDIMENSIONAL SCALING OF BINARY DATA FOR HOMOGENEOUS GROUPS

Robert P. [illegible]edinger

and

Jagdish N. Sheth

INTRODUCTION

Inspired by measurement in the hard sciences, the first developed techniques in multidimensional scaling (c.f., 20) required the input data to be metric. However, the necessity of using metric data as input required strong assumptions about the underlying psychological processes (9,11). One method of scaling psychological data while relaxing the assumptions of the input data and the concomitant cognitive processes is to collect lower order data (ordinal), find a function to transform this data into a metric representation, and then input this transformed data into existing metric multidimensional scaling techniques. Shepard (13,14) discusses the problems attendant with this approach and as an alternative presents a method of multidimensional scaling (refined by Kruskal (7,8)) that requires only ordinal data as input, yet produces scales with metric properties.

The major advantage of nonmetric versus metric multidimensional scaling is a relaxation in the assumptions of the underlying psychological processes an individual uses in making judgements. As Shepard (11) noted, qualitative judgements can be made with greater ease, assurance, validity, and reliability than can quantitative judgements. However, several problems can be identified with these nonmetric multidimensional scaling techniques.

First, an assumption of metric techniques is that the respondent be consistent throughout the task with respect to the criteria used and the

quantification of that criteria. Nonmetric techniques, while they do not require quantification, retain the assumption of consistency of criteria. Shepard (12) found that similarity judgements are likely to be influenced by attention fluctuations, and Torgerson (15) reported that the judgements may be affected by contextual effects.

Second, although the nonmetric methods require only ordinal properties in the data, the assumptions of ordinality must be met. If the basic ordinal properties (properties that are empirically testable) are exhibited by the data, the researcher is justified in using geometric models for scaling; thus, the use of nonmetric techniques depends on the validity of the underlying ordinal assumptions (1). The more difficult the task, the more likely it is that the underlying assumptions of the psychological process and of consistency will not be met.

Task difficulty can be resolved primarily as a function of the number of stimuli and the requirements of the task. As the number of stimuli increases, the difficulty of the task increases geometrically. The rank ordering of similarities of all possible pairs (990) of forty-five stimuli is a more difficult task than the rank ordering of all possible pairs (45) of ten stimuli. Rao and Katz (10) state that standard methods of collecting similarities data (for example, magnitude estimation, ranking of all possible pairs, or n-dimensional rank ordering) for large stimulus sets are cumbersome and may render judgements meaningless. Further, different techniques require different types of data, which affects task difficulty. The less invariant the data is to be (metric vs. ordinal), the more restrictive the assumptions of the underlying process, and hence the task will be more difficult. For example, the question "How much greater is A than B?", which would yield interval data, is a more difficult task than that represented by the question

"Which is greater, A or B?", which would yield ordinal data.

The third problem associated with nonmetric techniques is that these methods require assumptions on the part of the researcher as to the dimensionality of the underlying process and the metric to be used for calculating distances and scaling stimuli. The calculations in these techniques are based on the minimization of some criterion of error. Hence, if the underlying model (i.e., dimensionality and metric) is inappropriate, the procedures will calculate results capitalizing on the noise in the data, making interpretation difficult and statistical inferences to populations or across similar experiments unlikely (1).

What is needed then are simpler data collection procedures to handle the first two problems and simpler analytic procedures (at least in terms of fewest assumptions) to handle the third problem. Due to the large number of stimuli necessary for many marketing studies, attention has focused on providing alternative methods of collecting ordinal (similarities) data, methods which basically involve a reduction in the number of judgements the individual must make (10). However, an alternative solution is to reduce the difficulty of the task by further relaxing the assumptions underlying the psychological process implicit in the data collection technique. Rather than collecting ordinal data, the researcher can obtain nominal (classifactory) data or, in the simplest case of two classes, binary data. Green, Wind, and Jain (5) analysed associative data by assuming that the association frequency represented a proximity measure of the stimuli and utilized existing geometric scaling models to arrive at their configurations. They found that the technique resulted in high dimensionality which was difficult to interpret. They met the first condition of simpler data, but not the second condition of simpler analytic strategy which suggests that an alternative method of analysis fo associative data may also be appropriate.

The remainder of the paper describes a method of scaling associative (specifically binary) data which (1) requires as input only binary similarities thereby increasing the consistancy of the data while relaxing the assumptions of the underlying cognitive process, and (2) does not require prior specification of a geometric model (dimensionality and metric). After a discussion of the technique, the method is applied to the scaling of soft drinks and the results compared with the results from a standard multidimensional scaling method. Finally, the unresolved problems associated with this technique and the implications of the technique for marketing research are discussed.

DESCRIPTION OF THE MODEL

Binary data may be collected in a variety of ways, ultimately represented as the assignment of the stimuli to one of two groups. Judgements can be made regarding an object's possession of an attribute, or an object belonging to a group. To collect binary similarities data respondents would judge whether a pair of stimuli were similar or not similar. Accumulating judgements over individuals, a frequency distribution of the similarity of stimulus-pairs is obtained. Guttman (6) noted that a multivariate frequency distribution is scalable if one can derive from the distribution a quantitative variable with which to characterize the objects in the population so that each attribute is a simple function of that quantitative variable. Justified by the arguement that factor analysis can be legitimately applied to any symmetric table, Burt (3) describes a technique by which qualitative data can be factored analyzed. Sheth (10) has adapted this technique for the analysis of brand loyalty.

Suppose we wish to estimate the attribute space of n products and then scale the products within that space relying on binary similarities data for input. The similarity judgements are obtained by asking M individuals

whether a product-pair is similar (coded 1) or not similar (coded 0) for each of the $M = n(n-1)/2$ product-pairs. The data can be represented in an M x N matrix $\underline{Y}$, where each cell, $y_{i,k}$, represents the judgement of similarity of product-pair k by individual i.

$$\begin{array}{cc} & \text{product-pair } (k) \\ & \begin{array}{cccccc} 1 & 2 & \cdot & \cdot & \cdot & M \end{array} \\ \underline{Y} = & \left[\begin{array}{cccccc} 0 & 1 & \cdot & \cdot & \cdot & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{array}\right] \begin{array}{c} 1 \\ 2 \\ \cdot \\ \cdot \\ N \end{array} \quad \begin{array}{c} \text{individual} \\ (i) \end{array} \end{array}$$

In estimating the relevant attribute space, a necessary assumption is that all the individuals use the same space in making judgements. To test this assumption of homogeniety, a points of view analysis (22) using Eckart and Young's theorem of matrix approximation (4) is performed. An individual by individual matrix, $\underline{C}$, is calculated

$$\underline{C}_{M \times M} = \underline{Y}_{N \times M}\, \underline{Y}'_{M \times N}$$

where each cell, $c_{i,j}$, represents the number of times individuals i and j both rated a product-pair as similar. $\underline{C}$ turns out to be nothing more than a square symmetric contingency table. These absolute joint frequencies are a function of the number of product-pairs rated. To eliminate this sample size bias, the frequencies are standardized by computing the relative joint frequencies, $p_{i,j} = c_{i,j} / M$. Dividing these relative joint frequencies by the standard deviation, $(p_i p_j)^{\frac{1}{2}}$, results in a set of proportionate values

$$c^*_{i,j} = p_{i,j} / (p_i p_j)^{\frac{1}{2}} = c_{i,j} / (c_i c_j)^{\frac{1}{2}} \quad .$$

This is equivalent to pre- and post-multiplying $\underline{C}$, the contingency table, by a diagonal matrix $\underline{D}^{-\frac{1}{2}}$ with elements $1 / c_j^{\frac{1}{2}}$. Thus, we obtain a square symmetric matrix $\underline{R}$, which is positive, semi-definite;

$$\underline{R} = \underline{D}^{-\frac{1}{2}} \underline{C} \underline{D}^{-\frac{1}{2}} = \underline{D}^{-\frac{1}{2}} \underline{Y} \underline{Y}' \underline{D}^{-\frac{1}{2}} = \underline{M} \underline{M}' \quad \text{where } \underline{M} = \underline{D}^{-\frac{1}{2}} \underline{Y}$$

and being symmetric, $\underline{R}$ has grammian properties (2,17). This standardization yields 1's in the diagonal, hence $\underline{R}$ may be directly applied to principal components analysis, resulting in each individual R_j being expressed as a linear combination of factor scores, $\underline{F}$.

$$R_j = a_{j,1} F_1 + a_{j,2} F_2 + \ldots + a_{j,m} F_m$$

Using the factor scores, groups of individuals with assumed similar psychological attribute spaces can be formed. The subsequent scaling of products within an attribute space should be applied separately to each homogeneous group thus identified.

Scaling by Factor Analysis

Summing over individuals, a product by product square symmetric contingency table $\underline{X}$ is created for the group. Again, to eliminate sample size bias, $\underline{X}$ is standardized by calculating relative frequencies and dividing by the standard deviations.

$$X^*_{i,j} = X_{i,j} / (X_i X_j)^{\frac{1}{2}}$$

This standardized matrix, $\underline{X}^*$, is positive, semi-definite, and being symmetrical has grammian properties. Since the standardization yields 1's in the main diagonal, the matrix may be used directly in principal components analysis. $\underline{X}^*$ may be directly factored into the product of principal components $\underline{U}$ and

a matrix of characteristic roots, Λ^2, in the following manner. Since X^* is grammian, a matrix M can be found such that $X^* = M M'$. Defining U and W as transformation matricies such that $U = U^{-1}$ and $W = W^{-1}$, let $M = U \Lambda W$. Then,

$$X^* = M M' = (U \Lambda W) (W' \Lambda U') = U \Lambda^2 U' .$$

Each variable, X_j^*, can then be expressed as a linear combination of scores on the principal components, F, and the product-moment correlations, A, between the factors and the variables.

$$X^* = A F ; \text{ where, } A = U \Lambda , \text{ and } F = \Lambda^{-1} U' X^* .$$

The resulting principal component vectors, which are orthogonal, represent the underlying dimensions in the psychological process. Because the results are unique only up to affine transformations, the principal component vectors may be rotated to aid in identification. That is, a square symmetric matrix, T, with the restriction that $T T' = I$ can be found such that

$$X^* = U T \Lambda^2 T' U' = A_r F_r , \text{ where } A_r = U T \Lambda .$$

Factor scores for each product on the underlying dimension can be calculated using either the rotated solution

$$F_r = (T' A_r' A_r T)^{-1} T' A_r' X^*$$

or the unrotated solution[1]

$$F = (A' A)^{-1} A' X^* .$$

---

[1] This calculation is derived from the relationships $A = U \Lambda$ and $F = \Lambda^{-1} U' X^*$ as follows:

$$A \Lambda^{-1} = U ; \quad A' A \Lambda^{-1} = A' U ; \text{ since } (A' A) \text{ is invertible,}$$

$\Lambda^{-1} = (A' A)^{-1} A' U$; thus, $F = (A' A)^{-1} A' U U' X^* = (A' A)^{-1} A' X^*$ because $U U' = I$. Similarly for the rotated factors.

The factor scores represent the desired scale values of each of the products on the underlying dimensions, and a geometrical representation can be obtained from a plot of these scores. If the factor scores are computed after rotation, the rotation must be non-distance destroying or the resultant scale values will be meaningless.

There are several advantages of this method of scaling over the more traditional nonmetric multidimensional algorithms. First, this technique is not based on a criterion of error. Whereas geometric models attempt to best fit the data, that is to find a solution with interpoint distances whose rank order most closely approximates the rank order of the original data, this factor analytic technique attempts to explain the maximum amount of variation in the data. Second, in multidimensional scaling algorithms, the resultant scales on any dimension are dependent on the number of dimensions specified. However, the scale value of an item on a factor is independent of the number of factors specified because the factors are extracted sequentially in order of the amount of variation explained and are orthogonal. Finally, traditional methods can obtain a local minimum. That is, the techniques are dependent on the initial configuration specified by the researcher, even if it is only a random placement. Factor analysis requires no such initial starting point.

AN APPLICATION

The products used for this experiment were fifteen soft drinks: Coke, Pepsi, Royal Crown, Dr Pepper, Tab, Diet Pepsi, Seven-up, Sprite, Squirt, Diet Seven-up, Root Beer, Grape, Cherry, Orange, and Lemon-lime. Soft drinks were chosen because of subject experience with the product class, recognition by name only was possible, and the set of all possible soft drinks with which subjects were familiar was large.

A total of seventy subjects was divided into two equal groups. The first group was presented a list of all possible pairs (105) of the fifteen soft drinks and asked to indicate whether or not they considered the pair to be similar or not. The responses (yes for similar, no for not similar) constituted the binary data. One month later, each member of the group was presented a deck of cards, each card containing a pair of soft drinks. The subjects were asked to rank order the cards so that the top card was the pair judged most similar, the second card the next most similar and so forth. It was further suggested that the subjects use a stepwise procedure to complete the task, first sorting the cards into two piles of similar and dissimilar pairs, each pile then sorted into two piles, and so forth. After eight piles or so had been created, they were to rank the cards in each pile, combine piles one at a time, check the ordering of the new complete pile, and when completed, go through the deck one last time to be certain they were satisfied with the ordering. At the end of the task, the subjects were asked to describe the process and the criteria they used in completing the task, the perceived difficulty of the task, and their confidence in being able to replicate the process consistently. Several subjects were then given a second deck and asked to perform the task again as a measure of reliability. A similar procedure was used with the second group, except the order of the tasks was reversed (i.e., ranked data were collected first). Thus, for each individual, two sets of data were collected, namely a product by product matrix of binary similarities data and a product by product matrix of rank order similarities data.

Although both techniques required judgements on 105 pairs of products, the binary data task took less than one-fourth the time to complete than the rank order task. Further, the rank-order task was perceived as far more difficult than the binary task. While alternative methods of collecting rank order data are available, this stepwise method was chosen so that the results would be as

"accurate" as possible. Finally, the indepth questioning concerning the rank order task indicated that they were not consistent in their use of criteria for judging the similarities.

RESULTS AND DISCUSSION

Points of view analysis was performed on both sets of data, and in both instances, only one group appeared with no outliers. If more than one group had appeared, then separate scaling would have been performed for each subgroup. In this instance, all individuals were included in each analysis. Further, the data were analyzed separately for each group to determine if the order of the tasks had any effect on the results. There appeared to be no order effect, based on visual comparison of the resulting maps. Therefore, the two groups were combined and an analysis using the total sample was performed. Because of the high degree of homogeneity between the two groups, only the results from the analysis of the total sample is presented.

The Rank Order Data. A group similarities matrix was calculated with cell entries consisting of the average rank order for that product-pair; this matrix was used as input for TORSCA with the three dimensional results presented in Figure 1. As previously mentioned, this technique requires the prior specification of a model (metric and dimensionality) and of an initial configuration. For this study, the Euclidean distance function was chosen and 2-, 3-, 4-, and 5-dimensional solutions calculated, each starting from a random initial configuration. The scale values of a solution are dependent on the number of dimensions, hence, a necessary task for the researcher in applying these techniques is to choose the number of dimensions. A possible approach is to choose the dimensionality based on interpretability and the information provided (15). Stress values, measuring the goodness of fit of the data, can

be used. Stress values for the 2-, 3-, and 4-dimensional configurations were .240, .160, and .107 respectively. Primarily for the purpose of comparison with the binary data solutions, the three dimensional solution is presented.

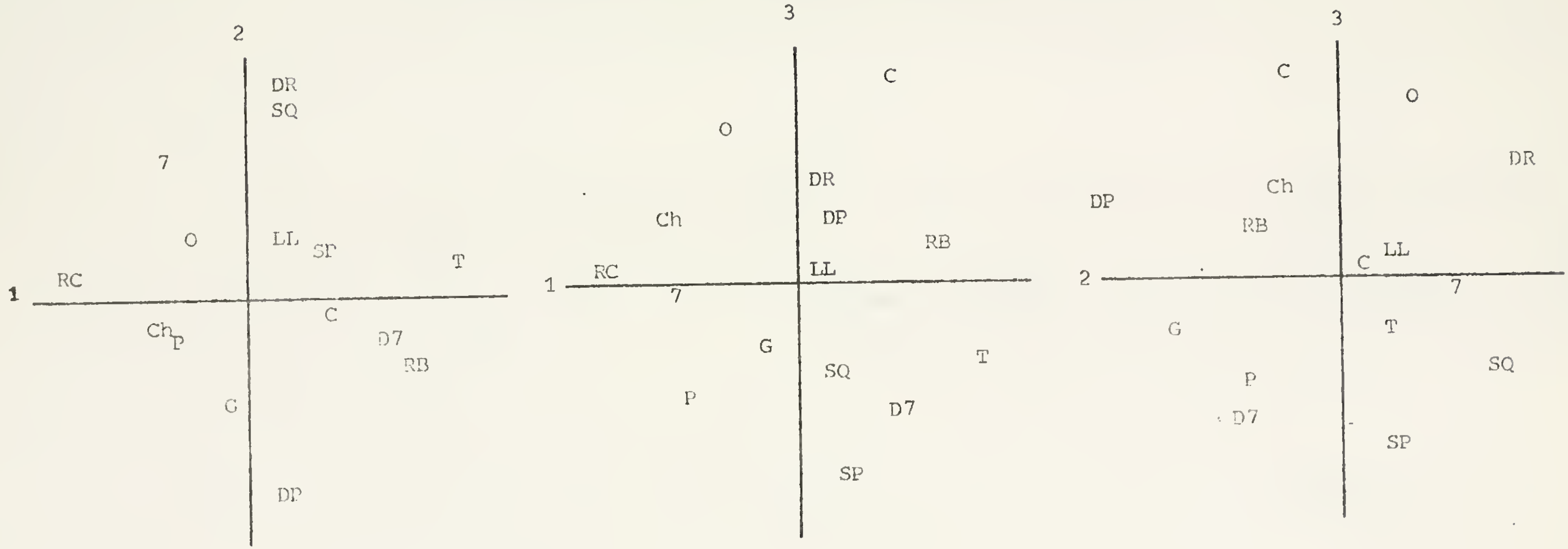
As is apparant from an examination of Figure 1, there is no easy and obvious interpretation of the results. This further demonstrates a problem with geometric models, namely interpretation of the results. Several possible methods to aid in the identification process include factor analyzing the data and using the factor loadings, or collecting evaluations of each product on various prespecified criteria and then fitting regression lines using this data to the obtained perceptual space. Also of interest in this example is that although the stress value decreased for the 4-, and 5-dimensional solutions, interpretation was not enhanced by the addition of the extra dimensions. This leads us to conclude that the underlying model implicit in the technique may not be appropriate. The recourse for the researcher is to continue to try additional models in the hope of obtaining a meaningful solution.

The Binary Data. The method of analysis described in this paper was applied to the group contingency table. Using a criterion of either significant eigenvalues or of common versus unique factors, the factor analytic procedure yielded a solution with three factors explaining slightly more than 80% of the variance. The plots of the rotated factor scores appear in Figure 2.

As opposed to the rank-order solutions, interpretation of these dimensions seems relatively apparant. The first dimension appears to be a cola (alternatively a dark-colored) dimension with seven products--Coke, Pepsi, Royal Crown, Tab, Diet Pepsi, Dr Pepper, and Root Beer--loading heavily. (Note, interpretation is aided in this technique by the use of the factor loadings). The second dimension appears to be an "un-cola" dimension (a lemon-lime, citrus flavored dimension) with five products--Seven-up, Sprite, Squirt, Diet Seven-up, and Lemon-lime--loading heavily. The third dimension appears

FIGURE 1

RANK ORDER DATA PRODUCT SPACE

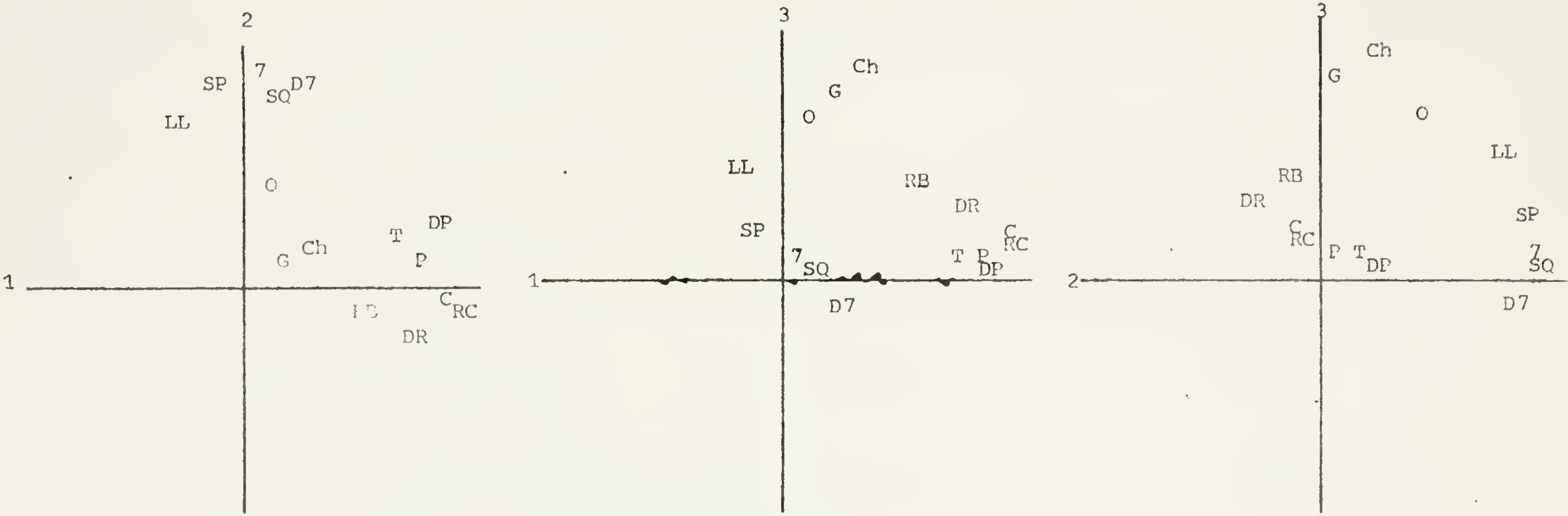


C - Coke
P - Pepsi
RC - Royal Crown
DR - Dr Pepper
T - Tab
DP - Diet Pepsi
RB - Root Beer

SQ - Squirt
SP - Sprite
7 - Seven-up
D7 - Diet Seven-up
O - Orange
Ch - Cherry
LL - Lemon-lime

FIGURE 2

BINARY DATA PRODUCT SPACE



C - Coke	SQ - Squirt
P - Pepsi	SP - Sprite
RC - Royal Crown	7 - Seven-up
DR - Dr Pepper	D7 - Diet Seven-up
T - Tab	O - Orange
DP - Diet Pepsi	Ch - Cherry
RB - Root Beer	LL - Lemon-lime
G - Grape	

TABLE I

MEAN AND RANGE OF RANK ORDER DATA

Drinks		Average Rank Order	Range of Rank Orders	Drinks		Average Rank Order	Range of Rank Orders	Drinks		Average Rank Order	Range of Rank Orders
2	1	49.93	19 - 103	3	2	49.93	7 - 101	3	1	58.73	15 - 103
[illegible]	3	47.[illegible]	[illegible]	4	2	50.70	1[illegible] - 102	4	1	60.3[illegible]	[illegible] - 99
[illegible]	4	[illegible]	25 - 105	5	3	62.27	14 - 10[illegible]	5	2	65.6[illegible]	2[illegible] - 105
[illegible]	1	67.73	23 - 104	5	5	70.47	40 - 105	[illegible]	4	65.[illegible]	24 - 100
6	5	62.[illegible]	26 - 96	6	2	55.50	16 - 102	[illegible]	1	18.20	2 - 70
[illegible]	4	20.[illegible]	4 - 69	7	5	70.00	12 - 103	7	4	67.47	8 - 103
[illegible]	3	58.[illegible]	6 - 93	7	2	54.97	5 - 104	7	1	11.33	1 - 40
[illegible]	[illegible]	17.[illegible]	3 - 42	8	5	31.97	6 - 87	8	5	73.10	20 - 101
[illegible]	4	64.[illegible]	27 - 105	8	3	60.63	31 - 92	8	2	56.07	16 - 105
8	1	20.[illegible]3	1 - 59	9	8	17.00	1 - 49	9	7	9.40	[illegible] - 31
9	6	10.[illegible]	1 - 69	9	5	69.07	33 - 102	9	4	63.55	21 - 103
9	3	58.[illegible]	2[illegible] - 94	9	2	56.00	13 - 102	9	1	10.73	1 - 32
10	9	66.37	10 - 105	10	9	75.43	29 - 105	10	7	71.17	29 - 101
10	6	55.03	6 - 105	10	5	49.10	16 - 95	10	4	64.30	29 - 104
[illegible]	3	63.03	21 - 102	10	2	75.87	35 - 103	10	1	70.87	28 - 105
11	10	10.20	1 - 89	11	9	70.17	10 - 102	11	8	74.90	29 - 105
11	7	62.77	9 - 102	11	6	52.80	4 - 104	11	5	42.30	12 - 104
11	4	55.96	24 - 105	11	3	58.67	9 - 99	11	2	72.10	23 - 103
11	[illegible]	74.93	24 - 104	12	11	45.07	14 - 95	12	10	45.27	15 - 96
12	9	66.40	9 - 105	12	8	72.63	11 - 104	12	7	71.03	24 - 105
12	6	72.70	33 - 104	12	5	43.60	7 - 100	12	4	58.97	27 - 105
12	3	49.03	4 - 102	12	2	71.20	37 - 103	12	1	71.63	10 - 103
13	12	36.80	6 - 94	13	11	22.77	5 - 101	13	10	19.37	4 - 86
13	9	63.43	13 - 103	13	8	73.73	25 - 105	13	7	64.97	22 - 102
13	6	66.57	19 - 102	13	5	41.47	9 - 105	13	4	64.33	37 - 102
13	3	60.20	18 - 94	13	2	73.60	45 - 103	13	1	69.37	14 - 101
14	13	7.37	1 - 26	14	12	33.73	2 - 93	14	11	18.67	9 - 80
14	10	13.03	1 - 79	14	9	64.90	7 - 105	14	8	70.33	15 - 105
14	7	63.13	9 - 104	14	6	58.57	22 - 103	14	5	39.90	4 - 103
14	4	65.47	29 - 104	14	3	55.50	17 - 105	14	2	70.23	42 - 102
14	1	62.47	25 - 105	15	14	[illegible].87	1 - 93	15	13	12.07	1 - 90
15	12	38.40	3 - 94	15	11	20.37	3 - 99	15	10	18.03	2 - 92
15	9	62.60	[illegible] - 105	15	8	74.43	26 - 105	15	7	68.10	21 - 104
15	6	70.40	20 - 103	15	5	38.97	4 - 102	15	4	62.63	26 - 105
15	3	53.77	3 - 97	15	2	73.03	43 - 105	15	1	74.47	30 - 104

to be a fruit-flavored (other than lemon-lime) dimension with three products--Cherry, Grape, and Orange--loading heavily and two products--Root Beer and Lemon-lime--loading slightly. Although not instructed to do so, the subjects seem to have used flavor as a major criteria in judging similarities resulting in three underlying flavor dimensions. Using the factor analytic technique, examination of alternative dimensional solutions is very easy because the factors are extracted independently of each other. In this instance, the fourth dimension in the four-dimensional solution (the other dimensions remaining basically the same) appeared that it might be a diet dimension with Diet Pepsi loading mainly heavily and Tab and Diet Seven-up loading slightly more that the rest (However, this factor could also be interpreted as a unique factor for diet Pepsi), but because of the criterion used for choosing dimensionality, was not included in the solution.

DISCUSSION

The purpose of both scaling techniques is to obtain a geometrical representation of the psychological space of soft drinks. In this study three dimensions were chosen for both techniques (1) for the purpose of comparison, (2) because these dimensions suited the criteria used in each technique, and (3) because a priori, three dimensions seemed appropriate (although not the same dimensions obtained in the binary solution). As is evident from a quick examination of Figures 1 and 2, the two methods did not yield similar results. Consequently, it is desirable to explain why these differences exist and to determine which mapping, if either, more nearly represents the true psychological space. We believe that the map resulting from the binary data provides a closer representation of the psychological space while the map from the rank order data is relatively meaningless. This belief is substantiated through the examination

of several comparative criteria of validity: cross validity, face validity, external validity, and predictive validity.

(1) The criterion of cross validity implies consistency of results across replications or across subgroups of the same population. In this instance, two separate sets of data applicable to each technique were originally collected. When analysed separately, the binary data yielded almost identical three dimensional perceptual maps for the two groups. However, the maps derived from the two sets of rank order data, while similar in the amount of dispersion exhibited, were completely different with respect to the relationships (interpoint distances) between the products.

(2) Results of a study have face validity if on inspection they are similar to what one might expect them to be. A priori, we hypothesized that the psychological space would be represented by three dimensions: a cola (color) dimension with Coke and Seven-up at the opposite ends, a diet dimension, and a fruit-flavored dimension. As noted, the map from the rank-order data was not interpretable, thus having no face validity. On the other hand, the binary data resulted in a map very nearly representing our a priori picture of the space. If we would have hypothesized that the cola-lemon dimension as being actually two orthogonal dimensions, then the four-dimensional binary solution--interpreting the fourth dimension as a diet dimension--would have almost exactly duplicated our a priori notions.

(3) As a measure of external validity, each subject was asked, after completing the rank order task, to state the criteria used during that task in the judgements of similarity. The most frequently mentioned criteria were cola, lemon flavor (seven-up like), diet, and fruit flavored. The technique using the binary data clearly extracted these dimensions, thus reproducing the stated criteria of the subjects. However, the rank order maps failed to even come close to these stated dimensions, eventhough these

external elicitations occurred immediately after the subjects performed the rank order task.

(4) Finally, predictive validity of the model can be obtained by having subjects produce geometric product spaces. Subjects were asked to physically place the products in three dimensional product spaces. Again, most subjects' maps were very nearly the same as those obtained by the binary scaling method. The only exceptions were a few subjects whose mapes were more nearly similar to our a priori dimensions of cola, fruit flavor, and diet.

The question then arises as to why a method utilizing weaker data (binary) produced results which across a variety of criteria were judged superior to those resulting from a technique utilizing stronger (ordinal) data. The first reason could be due to the different analytic procedures of the two methods. The traditional multidimensional scaling technique required prior specification of the model, and in this instance, our specification may have been incorrect, thereby yielding meaningless results. Further, the obtained results may have been one of several local minimums, dependent on the prespecified initial configuration. The factor analytic technique has neither of these problems since it requires no prior specification of a model or a starting configuration. Furthermore, the binary data technique produces dimensions that are extracted sequentially, thus facilitating the determination of the number of dimensions in the underlying perceptual space.

A second reason for the superiority of the binary data may be due to the differences in the data collection techniques. Both methods require consistency of criteria by the individual throughout the task and both must be applied to homogeneous groups of individuals. Thus, both methods must have data that are highly consistent, both within and between individuals. In the collection of the binary data, the task was rather simple. Subjects were able to complete the task in about ten minutes and afterwards indicated that they were able to

use the same criteria in making judgements throughout the task. Further, all subjects generally used the same criteria or at least judged the same pairs to be similar most of the time. In contrast, the rank order task was very difficult. On the average, the task required forty-five minutes to complete and all the subjects stated that they might have changed criteria during the course of the task. Subjects further indicated that they did not believe they would be consistent over trials, a fact we verified by repeat testing. Thus, the rank order task resulted in highly inconsistent data within subjects. To demonstrate the problem of between individual consistency, the average rank order (input to the TORSCA program) and the range of the rank orderings for each of the 105 product-pairs are presented in Table I. As can be seen from distribution of the rank orders, there is a tremendous discrepancy between individuals in the ranked (perceived) similarities of the product pairs (the smallest range is 26, the largest is 100). However, even if these between individual differences could be reduced, it is doubtful that meaningful results could be obtained from the rank order data because of the within individual inconsistency. The consistency problem results directly from the number of stimuli and the difficulty of the task.

CONCLUSION

Limitations of the Model

First, the technique presented is only applicable for group data. While some traditional multidimensional scaling techniques can be applied to similarities data for an individual as well as a group, this method requires relative frequencies as input, so the analysis can only be performed on group data. However, we are currently investigating a statistical procedure for mapping individual binary responses.

Second, this technique is only applicable in those instances where binary responses are appropriate. Other forms of associative data are not directly usable because of the need to form a frequency distribution of responses. Further, preference type data, often used in marketing applications of multidimensional scaling could not be scaled with this technique because the responses would not be binary.

Finally, this method also suffers from the problem of lack of invariance common in the nonmetric multidimensional scaling techniques. Since the results of this technique are unique only up to affine transformations, the axes chosen are somewhat arbitrary. Further, these is no exact criteria for chosing the number of dimensions. However, the criteria that do exist for this method are perhaps better substantiated than the criteria for other methods. Additionally, choice of the number of dimensions has little effect on the product positions on each dimension.

Summary and Implications

The implications for marketing research are many. The costs associated with binary data collection would be less. Less time is required per individual and compliance to cooperate in the task is higher; both should yield lower costs. Thus, even if binary data and ordinal data produced identical results, the use of binary techniques would be advantageous from a cost-benefit point of view.

Somewhat similar to cost effectiveness is the task effectiveness of this method. It is easier to maintain concentration for shorter tasks, all things else being equal. Further, considerably more information can be obtained in comparable time periods. Because the task difficulty is lower, within individual consistency will be higher.

Third, nonmetric methods are based on a criterion that minimizes some form of error, which results in a problem of statistical inference, especially if the underlying model (dimensionality and metric) is incorrect. The use of the frequency distributions of the binary data represents a method whereby statistical inference theory is applicable, which, through sampling, could result in generalizations to populations. In addition, if through points of view analysis, subgroups with different psychological spaces are found, statistical tests of differences between these subgroups are possible.

Finally, since marketing research typically involves large stimulus sets, if scaling is to provide useful analysis for the researcher, methodologies must be employed which have underlying assumptions that can be met. If the assumptions underlying a technique are not met, the validity of the results is questionable. Binary scaling represents one technique with assumptions that are more likely to be met, thus providing greater confidence in the validity of the results.

REFERENCES

1. Beals, Richard, David H. Krantz, and Amos Tversky, "Foundations of Multidimensional Scaling," Psychological Review, 75 (1968), 127-42.

2. Burt, Cyril. The Factors of the Mind. London: The University of London Press, 1940.

3. Burt, Cyril, "The Factor Analysis of Qualitative Data," The British Journal of Psychological, Statistical Section, 3 (1950), 166-85.

4. Eckart, Carl, and Gale Young, "The Approximation of One Matrix by Another of Lower Rank," Psychometrika, 1 (September 1936), 211-18.

5. Green, Paul E., Yoram Wind, and Arun K. Jain, "Analyzing Free-Response Data in Marketing Research," Journal of Marketing Research, X (February 1973), 45-52.

6. Guttman, Louis, "A Basis for Scaling Qualitative Data," American Sociological Review, 9 (1944), 139-50.

7. Kruskal, J. B., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, 29 (March 1964), 1-27.

8. ________________, "Nonmetric Multidimensional Scaling: A Numerical Method," Psychometrika, 29 (June 1964), 115-29.

9. Luce, R. D., and J. W. Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," Journal of Mathematical Psychology, 1 (1964), 1-27.

10. Rao, Vithala R., and Ralph Katz, "Alternative Multidimensional Scaling Methods for Large Stimulus Sets," Journal of Marketing Research, VIII (November 1971), 488-94.

11. Shepard, Roger N., "Metric Structures in Ordinal Data," Journal of Mathematical Psychology, 3 (1966), 287-315.

12. __________, "Attention and the Metric Structure of the Stimulus Space," Journal of Mathematical Psychology, 1 (1964), 54-87.

13. __________, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. I, "Psychometrika, 27 (1962), 125-40.

14. __________, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. II," Psychometrika, 27 (1962), 219-46.

15. __________, A. Kimball Romney, and Sara Beth Nerlove, Multidimensional Scaling: Theory and Applications in the Behavioral Sciences. New York: Seminar Press, 1972.

16. Sheth, Jagdish N., "A Factor Analytical Model of Brand Loyalty," Journal of Marketing Research, V (November 1968), 395-404.

17. __________, "Using Factor Analysis to Estimate Parameters," Journal of the American Statistical Association, 64 (September 1969), 808-22.

18. Torgerson, W. S., "Multidimensional Scaling of Similarity," Psychometrika, 30 (1965), 379-93.

19. __________, "Multidimensional Scaling: Theory and Method," Psychometrika, 17 (1952), 401-19.

20. __________, Theory and Methods of Scaling. New York: John Wiley & Sons, Inc., 1958.

21. Tucker, L. R., "Determination of Parameters of a Functional Relation by Factor Analysis," Psychometrika, 23 (1958), 19-23.

22. __________, and S. Messick, "An Individual Differences Model for Multidimensional Scaling," Psychometrika, 28 (1963), 333-67.

23. Young, Forest W., and W. S. Torgerson, "A FORTRAN IV Program for Nonmetric Multidimensional Scaling Analysis," Behavioral Science, 12 (November 1967), 498.